

Improvement ocean wave spectra estimation using the temporal structure of wave systems

Kpogo Nuwoklo Komlan Agbéko^{1,*}, Ailliot Pierre², Olagnon Michel¹, Guede Zakoua¹, Arnault Sabine³

¹ IFREMER, Ctr Brest, Lab Comportement Struct Mer, F-29280 Plouzane, France.

² Univ Bretagne Occidentale, Lab Mathemat Bretagne Atlant, Brest, France.

³ Univ Paris 06, LOCEAN IPSL, Paris, France.

Corresponding author : K. Agbéko Kpogo-Nuwoklo, email address : kkpogonu@ifremer.fr ; pierre.ailliot@univ-brest.fr ; michel.olagnon@ifremer.fr ; zakoua.guede@ifremer.fr ; sabine.arnault@locean-ipsl.upmc.fr

Abstract :

Sea states are usually the combination of several time-evolving wave systems whereas the classical spectral estimation methods assume stationarity. A method that adapts to the dynamical evolution of the spectral components is proposed to improve both omnidirectional and directional sea wave spectral estimations. In this method, periodograms are computed for each sea state as in the conventional methods, and rather than only smoothing individual periodograms, the overall time-history of periodograms are simultaneously smoothed in frequency and time dimensions. Since a simple two dimensional averaging would not be appropriate because the temporal evolution of the wave systems reflects typical non-stationary behaviors, we use either kriging or adaptive 2D kernel density estimators that allow the taking care of the spectral component frequency-time evolutions. The method is successfully validated on sequences of spectra typical of sea-state conditions in West Africa. The comparison with the simple 2D averaging method and individual periodogram smoothing method shows that the proposed method gives higher effective numbers of degrees of freedom, better estimates of the spectral shape and reliable spectral moments. The method also provides a tool for sea wave spectra interpolation and may thus be used to fill in missing values and improve wave systems tracking for storm identification purposes.

Highlights

► We propose a method to improve sea wave spectra estimation. ► The proposed method takes advantage of the structure of wave systems time-histories. ► The method also provides for missing spectra interpolation. ► We also apply the method to directional spectra estimation. ► The method is validated on spectra measured off Angola.

Keywords : Directional wave spectrum, Spectral estimation, Periodogram smoothing, Kriging

1 Introduction

Description of random ocean waves is of vital importance for many activities and therefore a considerable effort is spent on the collection and analysis of wave data. The most frequent approach for analyzing the characteristics of random ocean waves is spectral analysis [15], [4]. In that approach, a sea state is defined for a suitable short-term period of time (from half hour to several hours), and is considered fully characterized by its directional spectrum which provides the distribution of the wave energy as a function of frequency and direction. Reliable sea state spectral estimation is therefore important, especially in coastal engineering and others marine applications. Methods of two types are available for estimating the spectrum of a sea surface elevation time series: non-parametric and parametric methods. Non parametric sea wave spectral estimation generally consists in periodogram smoothing with frequency windows [26], [34], or in Fourier transformation of smoothed or truncated auto-covariance function [3]. In parametric spectral estimation, the task is to estimate the parameters of a model that describes the spectrum of the sea state [24], [18]. Parametric estimation is generally satisfactory for unimodal spectra, but much weaker when the spectrum exhibits many peaks [21].

In many cases, the sea state is a superimposition of coexisting wave systems (swells and wind sea) and the resulting spectra have many peaks. One often needs in such cases to partition the sea state into its wave systems for an accurate characterization (see [1], [11], [17], [25]) . Furthermore, some of successive wave systems are created by the same meteorological event (storm, hurricane, pressure low, etc.) and thus, one can track in time the evolution of wave systems in order to link them to the same event (see [17], [1]). Indeed, when looking at the time-history of periodograms, it is possible to observe temporal coherence in the spectral components. As an illustration, Figure 1

shows on the left panel a periodogram for individual sea state (in dark) and on the top right panel the time-history of periodograms off Angola. Linear patterns can be observed in the low frequencies ranges, corresponding to the time evolutions of swell systems. Some coherent patterns can also be observed in high frequencies ranges, corresponding to local winds effects. It is legitimate to expect that taking into account the temporal coherence of the spectral components may be a good way to improve sea wave spectra estimation. That allows to overcome the limits imposed to the estimation quality by strict stationnarity conditions.

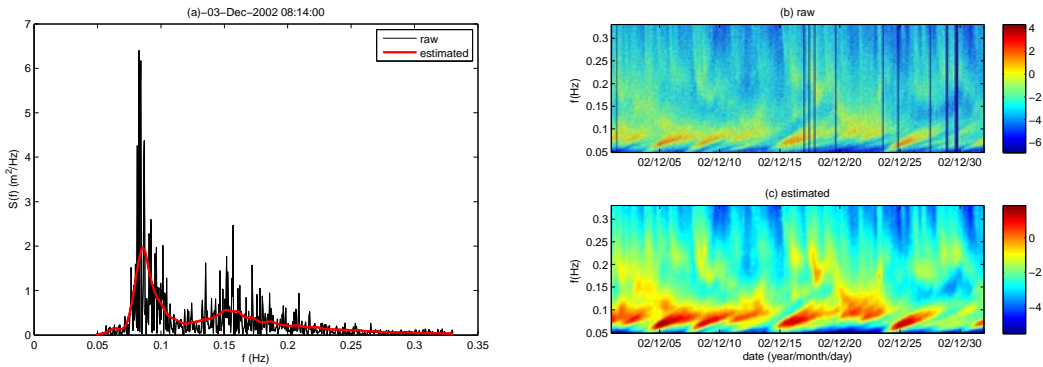


Fig. 1. Sea wave spectra estimation and interpolation applied to in-situ data: (left) example of periodogram and estimated spectrum; (right) time-history of periodogram (up) and estimated spectra (bellow).

This study proposes methods that adapt to the dynamical evolution of the spectral components to improve both omnidirectional and directional sea wave spectral estimation. Periodograms are first of all computed for each sea state as in the conventional methods. Yet, rather than only smoothing individual periodograms, we consider the overall time-history of periodograms as a random field that depends on frequency and time, and we smooth this field simultaneously in frequency and time dimensions. Since the temporal evolution of the wave systems reflects typical non-stationary behaviors, as for instance the linear increasing trend of the pic frequency of swell components, a simple two dimensional averaging would not be appropriate. Two techniques that allow to consider the dynamical evolution of wave systems are therefore proposed in this study to smooth the time-history of periodograms:

- Kriging that is a spatial optimal interpolation method based on the idea that the value at an unknown point should be a weighted average of the known values at its neighbors, where the optimal weights are calculated using the time and frequency correlations [29];
- Adaptive 2D kernel density estimator i.e. imposing a kernel that adapts to the dynamical evolution of wave systems [27].

The addition and the use of the time dimension, taking into consideration the dynamical evolution of the wave systems, improves the sea state spectral estimation. Spectra estimates using the proposed approach dramatically enhance the temporal tracking of wave systems for sea storms identification purposes. Moreover, in-situ wave measurements often suffer from missing data in consequence of instrument loss, malfunction, or delayed maintenance. For the many marine applications that need continuous measurements, the approach provides a straightforward way to

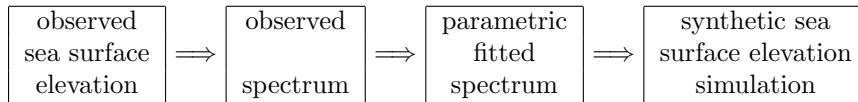


Table 1. Summary of the method used to generate synthetic time-series of sea surface elevation

interpolate spectra when faced with missing data (see Figure 1 right panel). The extension of the method to directional spectra improves also the estimation of the directional spreading which is important in some marine applications.

The paper is organized as follows. A brief description of the data used for this study is given in Section 2. In Section 3, we present the proposed method for omnidirectional spectral estimation. We then show some comparison results of the estimated spectra with those that are estimated by only smoothing individual periodograms, and with spectra that are estimated using a naive two dimensional averaging. The ability of the proposed method to interpolate missing spectra is finally shown. Section 4 presents the extension of the methods to the directional spectral estimation. Conclusions are given in Section 5

2 Data

The data used for this study are re-simulated times series of sea surface elevations derived from a time history of parametric spectra in order to validate the proposed methods in an idealized situation where the "true" spectra are known. The measurement data, covering period from March 2001 to April 2004 are obtained from a directional waverider buoy, sampled at 1.28 Hz, and anchored by 1450 m water depth off the coast of Angola. Half hourly successive spectra derived from in-situ measurements are first partitioned using a "dynamical partitioning algorithm" [2]. Synthetic spectra are then obtained as linear combinations of parametric models fitted to the partitions. A wind sea component is fitted by a Jonswap parametric model and swell components are fitted by Log-normal parametric models as recommended in [10], [21] for West Africa conditions (see Figure 2). The resulting time history of synthetic spectra provides a realistic sequence of spectra (reference spectra). From each synthetic spectrum, time series of sea surface elevation $z(\tau)$ are finally simulated, assuming that $z(\tau)$ is a Gaussian process, as described in [13], [22]. The goal is to estimate spectra from the re-simulated sea surface elevations times series with the proposed methods and the conventional ones and to compare the results with the reference spectra.

3 Proposed method for omnidirectional spectral estimation

3.1 Method

The method consists of two steps:

1. block the data by short stationary periods (typically 30 minutes) and compute periodograms for recorded sea elevation time series using discrete Fourier transform on each block;
2. smooth the time-history of log-periodogram using either:
 - kriging, or

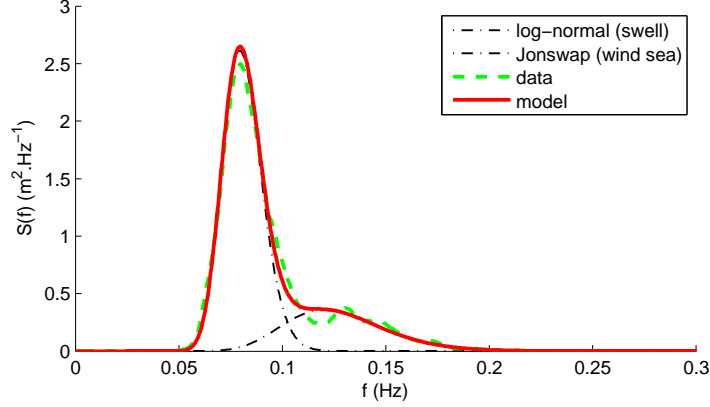


Fig. 2. Observed periodogram and fitted combined model (Log-normal function for the swell component and Jonswap for the wind sea).

- two dimensional Kernel Density Estimator by imposing a kernel that follows the dynamical evolution of waves systems (adaptive kde2D).

3.1.1 Definitions

Consider a real-valued, zero mean process $\{z_\tau\}_{0 \leq \tau \leq T}$ over duration T and assume that $\{z_\tau\}_{0 \leq \tau \leq T}$ is divided into N zero mean stationary sub-processes $\{z_\tau^1\}, \{z_\tau^2\}, \dots, \{z_\tau^N\}$ of individual duration $\delta t = T/N$ and with finite size observed realization $(z_0^1, \dots, z_{2n-1}^1), (z_0^2, \dots, z_{2n-1}^2), \dots, (z_0^N, \dots, z_{2n-1}^N)$ at time $t \in \{\delta t, 2\delta t, \dots, N\delta t\}$.

At a given time $t = j\delta t$, the associated periodogram for the stationary sub-process $\{z_\tau^j\}$ is computed for a discrete set of frequencies $f \in [\delta f, 2\delta f, \dots, n\delta f]$ as

$$I(f, t) = \frac{1}{2n} \left| \sum_{\tau=0}^{2n-1} z_\tau^j e^{-i2\pi f \tau} \right|^2 \quad (1)$$

with $\delta f = 1/(2n\tau)$ the sampling frequency.

Based on asymptotic properties, [6], [9] a frequently adopted model for $I(f, t)$ is

$$I(f, t) = S(f, t)\varphi(f, t)$$

with $\varphi(f, t)$ are $\overset{\text{iid}}{\sim} \exp(1)$ and $S(f, t)$ denotes the true spectral density. According to this model,

$$E[I(f, t)] = S(f, t)$$

and

$$Var[I(f, t)] = S^2(f, t).$$

Due to its very large variance, $I(f, t)$ is seldom used as an estimator of $S(f, t)$. A logarithmic transform is commonly applied to convert the above multiplicative noise model into an additive noise model and stabilize the variance (see, [9] [19] [23])

$$L(f, t) = m(f, t) + \xi(f, t)$$

where $L(f, t) = \log(I(f, t))$, $m(f, t) = \log(S(f, t))$ is the true log-spectral density and $\xi(f, t) = \log(\varphi(f, t))$ are *iid* random variables.

The periodogram estimate is an asymptotically (in n) unbiased but inconsistent estimate of the underlying true spectrum (see [30]). In particular, the estimate suffers from a high variance which does not converge to zero as n increases. Consistent estimators of wave spectral density can be obtained by assuming the spectrum constant over some neighborhood and directly smoothing the periodogram or the log-periodogram [32].

3.1.2 Smoothing the time-history of periodograms using Kriging

Kriging is based on the idea that the value at an unknown point should be a weighted average of the known values at its neighbors. Kriging theory provides an optimal linear unbiased estimator, based on assumptions on covariances and it is meant for estimation of a single realization of a random field [7].

We assume that the true log-spectral density $m(f, t)$ is a second order time stationary process and denote

$$\begin{cases} \mu_m(f) &= E[m(f, t)] \\ \text{cov}[m(f, t), m(f', t')] &= C_m(f, f', t' - t) \end{cases}$$

As a consequence the log-periodogram $L(f, t)$ is also a second order stationary process such that

$$\begin{aligned} \mu_L(f) &= E[L(f, t)] \\ &= E[m(f, t) + \xi(f, t)] \\ &= \mu_m(f) + \alpha \end{aligned} \tag{2}$$

with $\alpha = E[\xi(f, t)]$ and

$$\begin{aligned} \text{cov}[L(f, t), L(f', t')] &= \text{cov}[m(f, t) + \xi(f, t), m(f', t') + \xi(f', t')] \\ &= \begin{cases} C_m(f, f', t - t') & \text{if } (f, t) \neq (f', t') \\ C_m(f, f', t - t') + \sigma^2 & \text{if } (f, t) = (f', t') \end{cases} \end{aligned} \tag{3}$$

with $\sigma^2 = \text{var}[\xi(f, t)]$.

We construct an unbiased linear estimator of $m(f, t)$

$$\hat{m}(f, t) = \mu_m(f) + \sum_{\substack{i/ \\ |f_i - f| < Df \\ |t_i - t| < Dt}} \lambda_i (L(f_i, t_i) - \mu_L(f_i)) \tag{4}$$

with $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$ such that

$$E[(\hat{m}(f, t) - m(f, t))^2] = \lambda' K_L \lambda - 2\lambda' k_L + \text{var}[m(f, t)]$$

is minimum. λ is thus given by

$$\lambda = k_L K_L^{-1}$$

with

$$k_L = \begin{pmatrix} \text{cov}[m(f, t), L(f_1, t_1)] \\ \text{cov}[m(f, t), L(f_2, t_2)] \\ \vdots \\ \text{cov}[m(f, t), L(f_n, t_n)] \end{pmatrix} = \begin{pmatrix} C_m(f, f_1, t - t_1) \\ C_m(f, f_2, t - t_2) \\ \vdots \\ C_m(f, f_n, t - t_n) \end{pmatrix} \quad (5)$$

and

$$K_L = \begin{pmatrix} \text{var}[L(f_1, t_1)] & \text{cov}[L(f_1, t_1), L(f_2, t_2)] & \cdots & \text{cov}[L(f_1, t_1), L(f_n, t_n)] \\ \text{cov}[L(f_2, t_2), L(f_1, t_1)] & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}[L(f_n, t_n), L(f_1, t_1)] & \cdots & \cdots & \text{var}[L(f_n, t_n)] \end{pmatrix} \quad (6)$$

We estimate $\text{cov}[L(f, t), L(f', t + \Delta t)]$ by

$$\hat{C}_L(f, f', \Delta t) = \frac{1}{N} \sum_{t=1}^{N-\Delta t} L(f, t) L(f', t + \Delta t) - \frac{1}{N} \sum_{t=1}^{N-\Delta t} L(f, t) \sum_{t=1}^{N-\Delta t} L(f', t + \Delta t) \quad (7)$$

and $\mu_L(f)$ by

$$\hat{\mu}_L(f) = \frac{1}{N} \sum_{t=1}^N L(f, t)$$

and then deduce an estimate of μ_m and C_m using (2) and (3) and then the weights λ by replacing C_L and μ_L by the estimates in (5) and (6). Finally, the estimated spectral density at a given frequency and a given time is obtained by

$$\hat{S}(f, t) = \exp(\hat{m}(f, t)).$$

Note that it is also possible to compute the kriging variance

$$\text{var}[\hat{m}(f, t) - m(f, t)] = E[(\hat{m}(f, t) - m(f, t))^2] \quad (8)$$

which gives an estimation of the uncertainty and may be interesting for some applications.

In practice, the covariance is estimated by (7) using three year time-history of periodograms. In the case where only very short measurements durations would be available, covariance could still be estimated from long duration hindcast data. It is also possible to use a parametric model for the covariance function as it is often the case in geostatitics (see e.g. [7]). The covariance should be constant at a regional scale and one should be in position to use the same covariance to estimate spectra for any close enough site.

Figure 3 shows the correlation matrices defined as

$$R(f, f', \Delta t) = \frac{\hat{C}_L(f, f', \Delta t)}{\hat{C}_L(f, f', 0)}$$

A global yet expected observation is that the correlation decreases when Δt increases. Moreover, it is found that swell systems ($f \leq f_s = 0.11\text{Hz}$) are still self-correlated up to two days time-lags whereas wind sea systems are only weakly correlated beyond $24h$. For swell systems the correlation structure moves to the right side meaning that the correlation is more important between $S(f, t)$ and $S(f', t')$ when $t > t'$ and $f > f'$. This reflects the increasing trend of frequencies along a swell event (11). The correlation structure of wind sea systems is also found to move slightly to the opposite side reflecting the decreasing trend of wind sea frequencies. The parameters D_f and D_t of (4) which describe the size of neighborhood used for kriging interpolation are $D_f = 10\delta f$ and $D_t = 12\delta t$, where δf and δt denote frequency and time increments respectively.

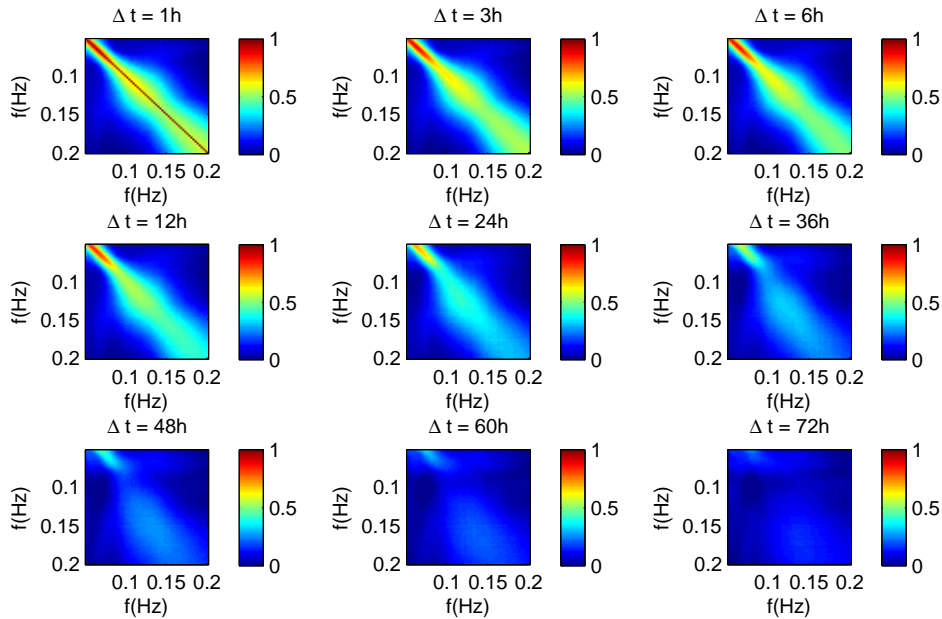


Fig. 3. Correlation function for different values of Δt

3.1.3 Smoothing the time-history of periodograms using 2D kernel density estimator

Kernel smoothing of the individual log-periodogram is a popular nonparametric method for spectral density estimation [16]. The usual kernel estimation (kde1D) for the log-spectrum is defined as

$$\hat{m}(f, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{f - f_i}{h}\right) L(f_i, t) \quad (9)$$

where kernel function K takes various forms in the literature, for instance a Hanning window, and is in fact a symmetric probability density function. The bandwidth h is a non negative smoothing parameter that controls the amount of smoothing.

Since the time- history of log-periodogram is assumed to be a random field $L(f, t)$, we propose in the following to smooth it with a 2D kernel density estimator. The general form for the 2D density estimator is then

$$\hat{m}(f, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\det(\mathbf{H})} K \left\{ \mathbf{H}^{-1} \left(\begin{bmatrix} f \\ t \end{bmatrix} - \begin{bmatrix} f_i \\ t_i \end{bmatrix} \right) \right\} L(f_i, t_i) \quad (10)$$

n is the number of observation points, K is the Kernel function, and \mathbf{H} is the 2×2 bandwidth matrix. When \mathbf{H} is a diagonal matrix i.e. $\mathbf{H} = \begin{bmatrix} h_f & 0 \\ 0 & h_t \end{bmatrix}$ where the elements h_f and h_t control the smoothing in frequency and time ranges respectively, the corresponding estimator is equivalent to a simple 2D averaging (see Figure 4 (b)) and it is denoted hereafter by "kde2D".

Yet we adapt in our proposed method a kernel that follows the frequency-time evolutions of wave systems. This is easily carried out by using a full bandwidth matrix i.e $\mathbf{H} = \begin{bmatrix} h_f & \lambda \\ \lambda & h_t \end{bmatrix}$ where λ is used to control the frequency-time evolution of wave systems ("adaptive kde2D"). According to a linear wave theory in deep water, for a wind event that was bounded in time and space, the dominant wave frequency observed at some distant location linearly increases over time at a rate inversely proportional to the distance traveled (see [12], [17]):

$$\frac{df}{dt} = \frac{g}{4\pi d} \quad (11)$$

The dynamical time evolution of wind sea systems however does not have any particular a priori pattern. It depends on the wind field characteristics (speed, direction, and time-displacement). Yet, according to Sverdrup and Munk [31], at the beginning of wind waves development, the energy transferred from wind to waves helps to increase the period (decrease the frequency). For simplicity it is assumed that the frequency-time evolution of wind sea systems follows in practice a decreasing linear trend (Figure 4 (c)). Thus λ has a positive value for swell systems and negative value for wind sea systems. The use of a full bandwidth matrix allows to take into account the dynamical evolution of the wave systems and therefore allows to overcome the limits imposed to the estimation by strict stationnarity conditions.

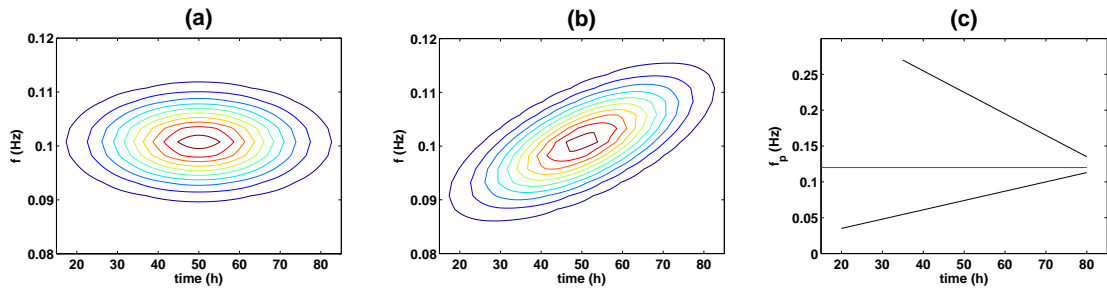


Fig. 4. Illustration of 2D Kernel Density Estimator: (a) with diagonal bandwidth matrix (kde2D), (b) with full bandwidth matrix (adaptive kde2D) for the case $f < f_s$; (c) Wave systems peak frequency time evolution

It is well known that the choice of \mathbf{H} is much more crucial than the choice of K (e.g., see [33], [28]). If the smoothing parameters are too small, the estimate is too noisy, exhibiting high

variability. If they are too large, then the estimate may miss key features due to over-smoothing, washing out small details. In practice, to speed up computation and for its simplicity we use a triangular kernel defined by

$$K(u) = \begin{cases} 1 - |u| & \text{if } |u| \leq 1. \\ 0 & \text{otherwise} \end{cases}$$

The bandwidth of the kernel is a free parameter that exhibits a strong influence on the resulting estimate. There are several methods to optimize the kernel bandwidth. In this study, the bandwidth is obtained by minimizing the mean integrated squared error (MISE) function. It is the most common optimality criterion used to select kernel bandwidth. The MISE is defined as

$$\text{MISE}(\mathbf{H}) = E \sum_f \sum_t (\hat{m}(f, t) - m(f, t))^2$$

As the theoretical log-density $m(f, t)$ is unknown, the MISE can not be estimated directly. Bowman [5] has proposed the use of cross validation method to estimate the MISE. In this method, the observation sample i is removed for estimating $m(f_i, t_i)$.

For instance we find for West Africa the optimal bandwidth matrix $\mathbf{H} = \begin{bmatrix} h_f & \lambda \\ \lambda & h_t \end{bmatrix}$ for the adaptive 2D kernel density estimator which elements are:

$$\begin{aligned} h_f &= 0.01Hz & h_t &= 15h & \lambda &= 0.45 \times mHz.h^{-1}, & (\text{case of swell systems}) \\ h_f &= 0.04Hz & h_t &= 9h & \lambda &= -1.2 \times mHz.h^{-1}, & (\text{case of wind sea systems}). \end{aligned}$$

The swell systems frequency domain smoothing bandwidth is smaller than the wind sea one justifying the fact that swells have narrower ranges of frequencies than wind sea; swell waves have dispersed from their generation area and have been dissipated. In contrast, the time domain smoothing bandwidth of swell systems is greater than the wind sea one because swells are generated by distant storms and the more remote is the storm, the greater is the duration between arrivals of successive frequencies according to dispersion of sea waves in deep water (11). Similarly to the covariance in the case of kriging, the bandwidth matrix should be constant at a regional scale. One needs to estimate in practice only once the bandwidth matrix which can be therefore used for other neighboring sites. Knowledge of the local climate would allow one to choose without computation bandwidth matrix values close to these optimal ones.

3.2 Results

Re-simulated time-series of sea surface elevation (see Section 2) are used to estimate spectra by the proposed methods. Periodograms are first computed from the time-series and the time-history of periodograms are smoothed using kriging technique of Section 3.1.2 and adaptive 2D kernel density estimator technique (adaptive kde2D) of Section 3.1.3. Our goal is to compare the estimated spectra with the reference spectra (i.e. synthetic spectra that are used to simulate the time-series of sea surface elevation). Furthermore, in order to check the relevance of our method, the spectra are also estimated, using a diagonal 2D kernel density estimator (kde2D) i.e. a naive 2D averaging, and using 1D kernel density estimator, that consists of periodogram smoothing in the frequency domain only (kde1D).

The first part of this section shows results related to the improvement in the spectral estimation using kriging and adaptive kde2D. The second part shows some results of the ability of the these techniques to interpolate missing spectra.

3.2.1 Improvement in spectral estimation

Figure 5 shows the time-history of the error $S(f, t) - \hat{S}(f, t)$, where $S(f, t)$ denotes the true spectra and $\hat{S}(f, t)$ the spectra estimated by the different methods (kde1D, kde2D, adaptive kde2D and kriging). The error is greater for the spectra estimated by kde1D and kde2D than those estimated using adaptive kde2D and kriging. In particular, kde1D and kde2D often underestimate the peaks of swells systems in comparison to the proposed methods. Visual inspection shows that spectra estimated using an adaptive kde2D or kriging are closer to the true spectra than those obtained by kde1D and kde2D. The same observation can be made on the individual spectra (see Figure 6).

The results are also studied quantitatively on three statistical parameters: effective number of degrees of freedom of the spectral estimation, Goda's peakedness parameter, and the spectral moments.

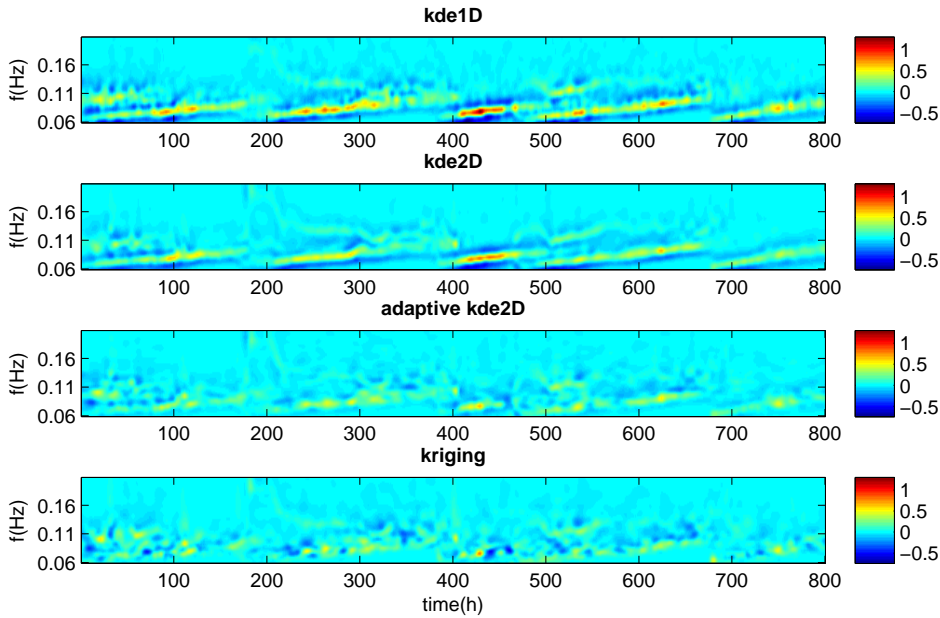


Fig. 5. time-history of the difference between the true spectra and the estimated spectra.

- Effective number of degrees of freedom

Under the assumption that the periodograms are independent and exponentially distributed, the spectrum $\hat{S}(f, t)$ obtained by smoothing the periodograms, is shown (see for example [8], [35]) to approximately obey a χ^2 distribution with $2\nu_{eff}$ degrees of freedom

$$2\nu_{eff} \frac{\hat{S}(f, t)}{S(f, t)} \sim \chi^2(2\nu_{eff}).$$

ν_{eff} is often used in spectral estimation as an indicator of goodness of the estimation. A classical assumption made to compute the number of degrees of freedom is that the spectrum

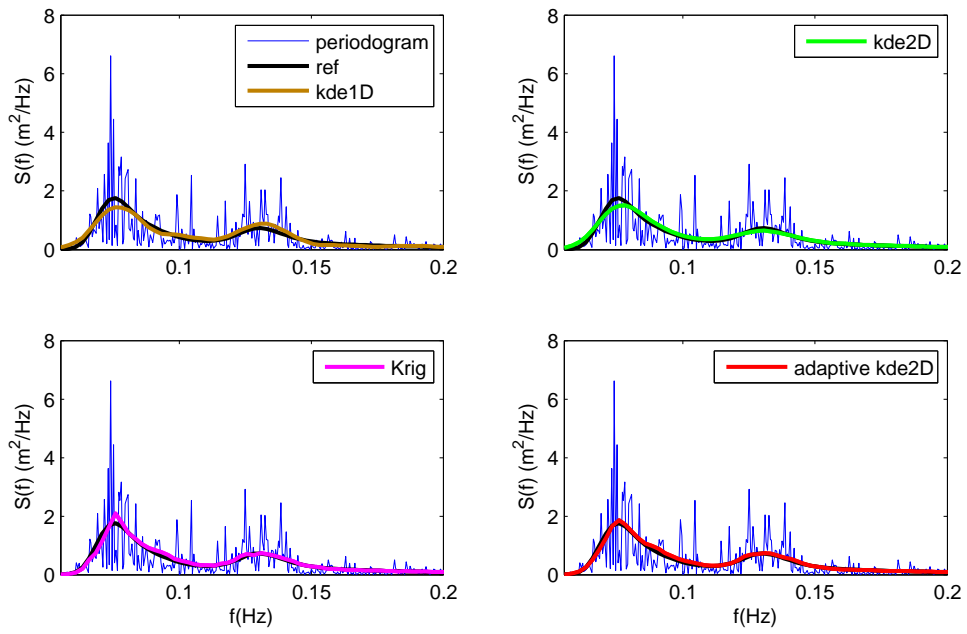


Fig. 6. Example of estimated spectrum in comparison with the reference one.

is constant in some neighborhood. Since this assumption is not fully valid, we compute an "effective number of degrees of freedom" defined by

$$\begin{aligned} \frac{1}{\nu_{eff}(f,t)} &= var \left[\frac{\hat{S}(f,t)}{S(f,t)} - 1 \right] \\ &= E \left[\left(\frac{\hat{S}(f,t)}{S(f,t)} - 1 \right)^2 \right], \end{aligned} \quad (12)$$

where $S(f,t)$ is the reference spectrum denoted by $S_{ref}(f,t)$ and $\hat{S}(f,t)$ is the estimated spectrum. The "effective number of degrees of freedom" integrated for the whole time-history can then be estimated by

$$\hat{\nu}_{eff}(f) = \frac{1}{\frac{1}{N} \sum_{t=1}^N \left(\frac{\hat{S}(f,t)}{S(f,t)} - 1 \right)^2} \quad (13)$$

where N is the total number of spectra. ν_{eff} can also be seen as the inverse of the normalized mean square error.

The Figure 7 shows the results of the "effective number of degrees of freedom" for the four estimation methods. Globally, $\nu_{eff}(f)$ has higher values for kriging and adaptive kde2D spectra than naive kde2D spectra which are slightly superior to the kde1D ones. $\nu_{eff}(f)$ varies proportionally with the spectral density and then reaches its peaks approximately at the same frequency as the peak frequency of swell and the wind sea systems. The peaks of swell are better estimated by kriging and by adaptive kde2D than by kde2D and kde1D. For wind sea systems the differences between adaptive kde2D, kde2D and kriging are not significant. The quality of the estimation is not good enough for the frequency interval between swell and wind sea systems ($[0.11 \text{ } 0.13\text{Hz}]$) and the tail of the spectrum $f > 0.18\text{Hz}$. For those intervals of frequency, the spectral densities are low and no temporal structure can actually be identified.

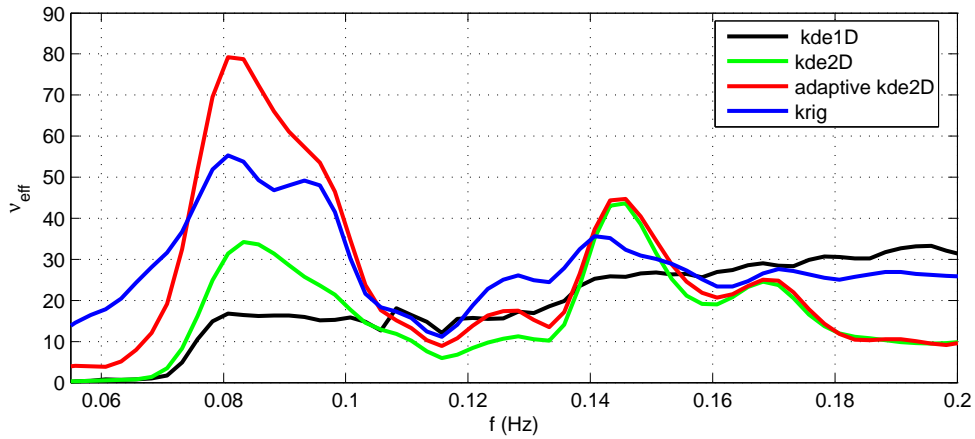


Fig. 7. Effective number of degrees of freedom as function of frequency for the three estimation methods

- Goda's peakedness parameter [14]

The objective is to investigate what improvements in the accuracy of the spectral bandwidth estimation one can have with our proposed methods. We used as spectral bandwidth the Goda's peakedness parameter Q_p , and it is estimated for each component (wind sea and swell) of the spectrum as:

$$Q_p = \frac{2}{m_0^2} \int_{f_{min}}^{f_{max}} f S(f)^2 df$$

where $m_0 = \int_{f_{min}}^{f_{max}} S(f) df$, f_{min} and f_{max} are respectively the lower and the upper frequencies bounds. We compute for the four estimation methods the normalized mean absolute error defined as

$$NMAE_{Q_p} = \frac{|Q_p - \hat{Q}_p|}{Q_p} \times 100$$

where Q_p is the reference Goda's peakedness parameter and \hat{Q}_p is the one that is computed from the spectra estimated by kde1D, kde2D, adaptive kde2D or kriging. The results Figure 8 shows the average of $NMAE_{Q_p}$ versus the peak frequency. The swell systems spectral bandwidths are clearly better estimated using adaptive kde2D and kriging than the others. Taking to account the frequency-time evolution of the swell systems allows to reduce approximately by 10% the error in the estimation of the spectral bandwidth. The improvement in the wind sea systems bandwidth parameter estimation is not significant. The wind sea have in fact a broad spectral bandwidth and short stationary durations and it is difficult to have precision in its estimation.

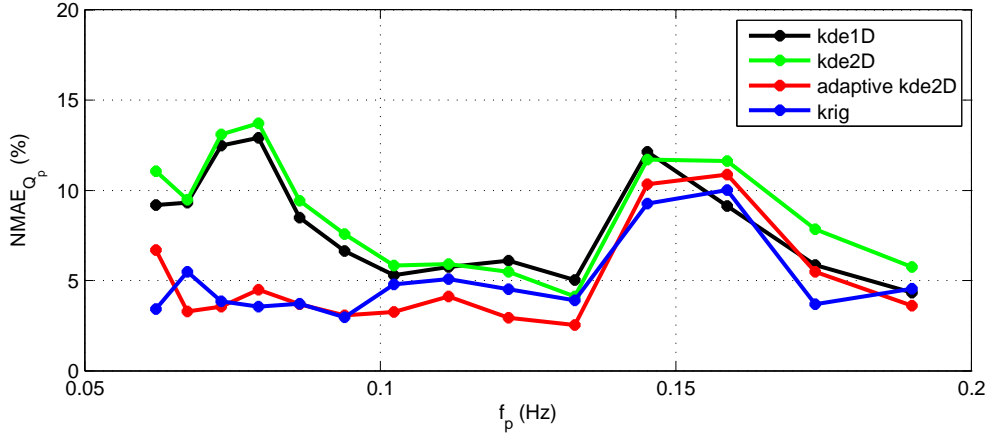


Fig. 8. Normalized mean absolute error for the Goda's peakedness parameter as function of the peak frequency

- Spectral moments

Spectral moments are used to compute sea wave parameters (significant wave height, mean period, etc). The ability of the proposed methods to give good estimates of spectral moments

is therefore another important criterion for choosing spectral estimates for ocean waves. The objective is to verify that the reduction of the variance in the spectral estimation using our method does not decrease the accuracy of the spectral moments. We thus investigate the relative difference between the reference spectral moments and the estimated ones defined for the n^{th} moment as:

$$Dm_n(t) = \frac{\left| \int f^n S_{ref}(f, t) df - \int f^n \hat{S}(f, t) df \right|}{\int f^n S_{ref}(f, t) df}$$

The Figure 9 shows the boxplot of Dm_n distribution for the $-1, 0, 1, 2$ orders of spectral moments. The difference between the reference spectral moments and the estimated ones is lower for kriging and adaptive kde2D. It can be concluded that the spectral moments, and thus the sea wave environmental parameters, may be accurately estimated by the two proposed methods. The reduction of the spectral variance by taking into account the temporal evolution of wave systems is thus very useful and does not affect the accuracy of the sea state parameters estimation.

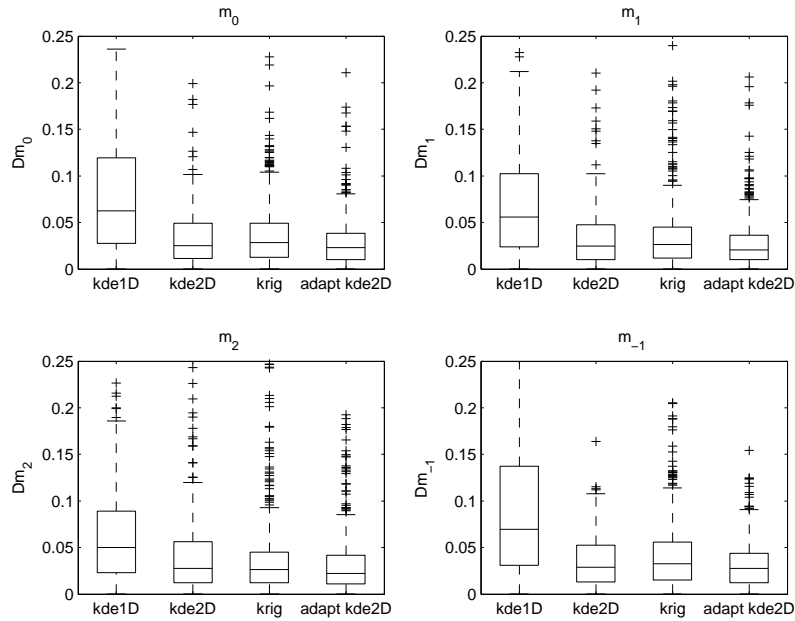


Fig. 9. Boxplot of relative error between the reference spectral moments and the estimated ones

3.2.2 Interpolation ability

In-situ deployed instruments, such as buoys, are sources for continuous long-term oceanographic data acquisition. Given the rough and unfavorable conditions in the ocean, data gaps are a common consequence of instrument loss, malfunctioning, or delayed maintenance or data collection, and are

a major problem for data analysis. The previously proposed methods, namely kriging and adaptive kde2D, can be effective tools for missing data interpolation. The ability of the methods to perform spectral interpolation is shown in the present part.

We take a time-history of periodograms $I(f, t)_{1 \leq t \leq N}$ split at half-hour intervals. For each time step t , gaps with durations $dt \in \{1h, 3h, 6h, 12h, 24h\}$ are arbitrarily generated by removing the corresponding number of periodograms $\{2, 6, 12, 24, 48\}$ centered on t . Spectra in the gap extent are then interpolated with kriging and adaptive kde2D methods. We then compare the interpolated spectra with the reference ones.

The Figure 10 shows an interpolated spectrum for different gap lengths. From the visual inspection of this figure we can see a good agreement between the reference spectrum and the two interpolated ones onto 12h. The overall accuracy of the interpolated spectra as a function of frequency is evaluated by the "effective number of degrees of freedom" (13). The "effective number of degrees of freedom" curves are presented in Figure 11 for the two methods and for different gap lengths. Globally, as could have been expected ν_{eff} decreases when the length of the gaps increases, reflecting that the quality of the interpolation decreases when the length of the gaps increases. The first one (0h) correspond to ν_{eff} when there is no gap; it is the same as in Figure 7. Up to 6h, the kriging ν_{eff} curves are most of the time under the adaptive kde2D ones, reflecting that adaptive kde2D is slightly better than kriging for those gap lengths. Regarding then 12h gap length, kriging interpolation is slightly better than the adaptive kde2D. The available information for this gap length is not sufficiently rich for accurate estimation of the peaks that correspond to the highest densities values. Above a gap length of 24h, spectra can not be interpolated by adaptive kde2D since the optimal time domain bandwidth of wind sea systems h_{tws} is equivalent to 9h. Kriging interpolation for those gap lengths also shows poor quality. It is clear that longer periods of missing spectra can not be properly interpolated.

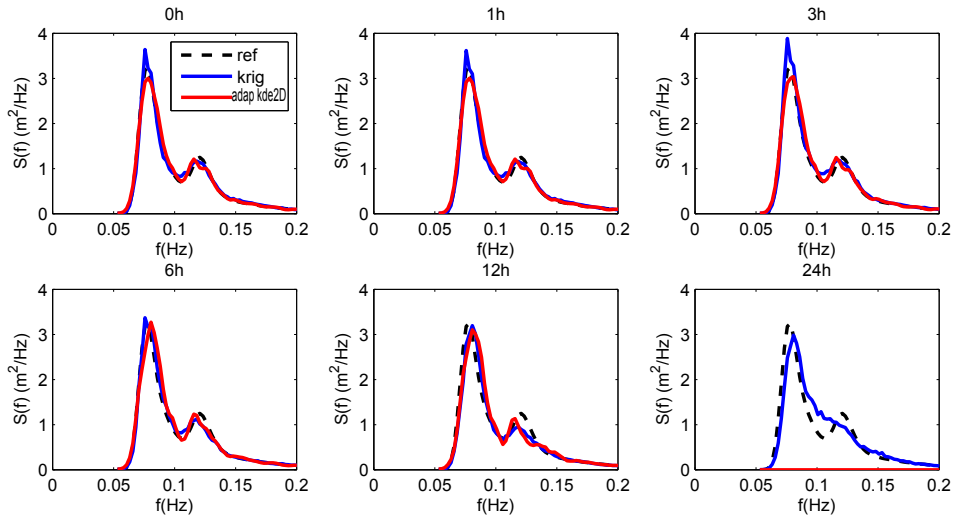


Fig. 10. Example of interpolated spectrum with the reference one for different gap length.

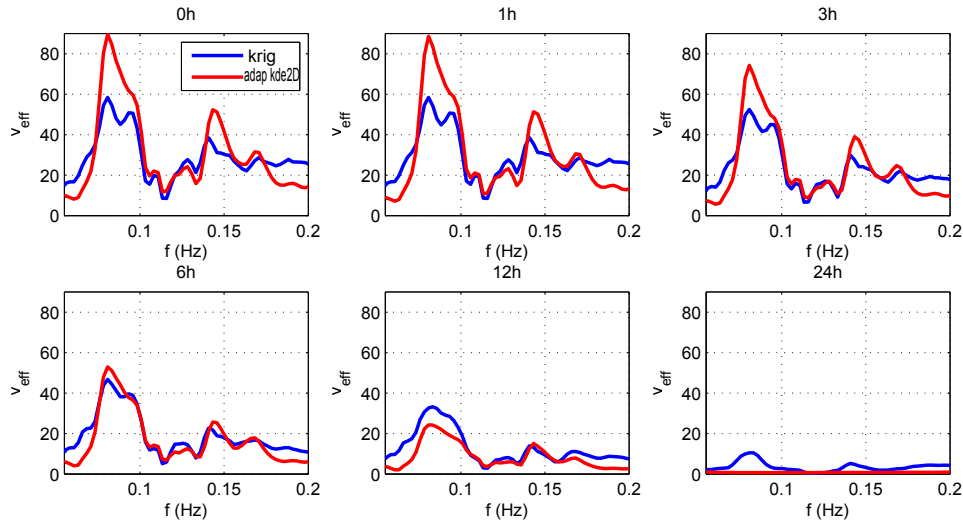


Fig. 11. "Effective number of degrees of freedom" ν_{eff} for the two methods and for different gap length.

3.2.3 Discussion

The results show that the spectral components temporal correlation structure can be used to improve sea wave spectrum estimation. Both the kriging method and the adaptive kde2D method give better estimates for spectral shape and for overall spectral moments than the kde1D method or naive kde2D method. The dependence between successive sea states inside the same meteorological event can not be taken into account when smoothing an individual sea wave periodogram. This dependence needs to be taken into account by also considering additional information provided by neighboring periodograms. However, as it is shown by the results, a simple 2D averaging (kde2D) of the periodograms is not sufficient. Introducing the time dimension in the spectral estimation requires to take care of the spectral components evolutions in order to have reliable spectra that can be used with confidence in marine applications.

Moreover, the results also show that between adaptive kde2D and kriging, the first has a slight superiority. They are roughly equivalent in the sense that we are looking in the two cases for a linear estimator of the spectral density $S(f, t)$ by using the neighboring spectral densities weighted by the neighbors distance to $S(f, t)$. The difference between the two methods is probably due to the size of data used to estimate the covariance function which is a key input in the kriging method. The latest results are obtained with covariance estimated from only three years data. Yet, even with three years data, the quality of spectra estimates using kriging is better than those obtained with kde1D or a simple 2D averaging. The accuracy of spectra estimates using kriging might be improved by using more than three years. Moreover, in contrast to adaptive kde2D, the kriging method has the advantage of not requiring a prior model for the time evolution of spectral components. The time-evolution of peak frequencies of swell systems is usually easy to model, but the of wind sea one lacks specific repeatable models. Choosing a prior model for wind sea peak frequency time-evolution may slightly affect spectral estimation using adaptive kde2D.

4 Extention of the directional spectral estimation

Directional information is crucial in many ocean engineering applications. This information is condensed in a directional wave spectrum described by

$$S(f, \theta) = S(f)D(f, \theta), \quad (14)$$

where $S(f)$ is the energy density spectrum and $D(f, \theta)$ is the directional spreading function. The directional spreading function can be expanded in an angular Fourier series,

$$D(f, \theta) = \frac{1}{\pi} \left[\frac{a_0(f)}{2} + \sum_{n=1}^{\infty} a_n(f) \cdot \cos(n \cdot \theta) + b_n(f) \cdot \sin(n \cdot \theta) \right], \quad (15)$$

where $a_n(f)$ and $b_n(f)$ are the Fourier coefficients. Based on three simultaneous wave measurements recorded at the same location, such as for instance heave, pitch and roll of a buoy, it is known that only the first and second Fourier order coefficients can be obtained based on the cross-spectra ([20]). It common to use a parametric form for $D(f, \theta)$, one of the most used directional spreading function model is *cosine - 2s* defined by

$$D(f, \theta) = K(f) \cos^{2s(f)} \left(\frac{\theta - \theta_m(f)}{2} \right), \quad (16)$$

where $K(f)$ is a normalization factor, $\theta_m(f)$ the mean wave direction, and $s(f)$ the directional spreading parameter.

In the case of the measurements recorded by a pitch/roll buoy the first Fourier coefficients may be obtained by

$$a_1(f) = \frac{Q_{zx}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, b_1(f) = \frac{Q_{zy}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, \quad (17)$$

where subscripts x , y , and z denote East, North, and vertical particle displacements respectively; and the real and imaginary parts C_{ij} and Q_{ij} of a cross-spectrum between wave records i and j . The mean wave direction, and the directional width are related to the first harmonic through

$$\theta_{m1}(f) = \arg(a_1(f) + ib_1(f))$$

$$\sigma(f) = \sqrt{2(1 - (a_1^2(f) + b_1^2(f)))}.$$

The directional spreading parameter $s(f)$ is then given by

$$s(f) = \frac{2}{\sigma^2(f)} - 1.$$

4.1 Method

The generalization of the methods to the 2D spectra estimation can be described by three steps.

1. Cross-periodograms computation

The periodogram of sea surface elevation $z(\tau)$ (1) can be seen as a cross-periodogram between z and itself. Generally the cross-periodogram between discrete-time signals $y(\tau)$ and $x(\tau)$ of length n is computed as

$$\begin{aligned} I_{xy}(f, t) &= \frac{1}{N} X^*(f, t) Y(f, t) \\ &= C_{xy}(f, t) - iQ_{xy}(f, t) \end{aligned} \quad (18)$$

with

$$X^*(f, t) = \frac{1}{2n} \sum_{\tau=0}^{2n-1} x_{\tau}^j e^{i2\pi f\tau}$$

and

$$Y(f, t) = \frac{1}{2n} \sum_{\tau=0}^{2n-1} y_{\tau}^j e^{-i2\pi f\tau}.$$

Following (18) the periodogram of the co-spectra ($C_{zz}(f, t)$, $C_{xx}(f, t)$, $C_{yy}(f, t)$) and the quadrature spectra ($Q_{zx}(f, t)$, $Q_{zy}(f, t)$, $C_{xy}(f, t)$) can be computed.

2. Smoothing the time-history of the logarithmic periodograms of the co-spectra and the quadrature spectra

The time-history of the logarithmic periodograms of the co-spectra and the quadrature spectra are treated as a random field and are separately smoothed using the same smoothing method (Kriging (Section 3.1.2) or kde2D (Section10)) as in the 1D case in order to get estimates of co-spectral densities ($\hat{C}_{zz}(f, t)$, $\hat{C}_{xx}(f, t)$, $\hat{C}_{yy}(f, t)$ and $\hat{C}_{xy}(f, t)$), and estimates of quadrature spectral densities ($\hat{Q}_{zx}(f, t)$ and $\hat{Q}_{zy}(f, t)$).

3. Directional spreading function

From those estimated cross-densities, the last step is to calculate the Fourier coefficients (17), then the directional spreading function $D(f, \theta)$ using (16) and finally the directional spectrum $S(f, \theta)$ using (14).

4.2 Results

As for the 1D case the validation is carried out on a simulated time series of sea surface horizontal ($x(\tau)$ and $y(\tau)$) and vertical ($z(\tau)$) motions. The simulated time series are derived from synthetic directional spectra obtained by fitting of parametric models on partitions derived from West Africa in-situ data. The wind sea component is fitted by a Jonswap parametric model, swell components are fitted by a Log-normal shape, and the directional spreading function is modeled by a \cos^{2s} function. To avoid imposing ill-estimated values of s for partitions of low energy, we use for at times the empirical directional width parameter σ derived in [10] for West Africa locations. σ is a function of frequency and can be expressed as:

$$\sigma(f) = 0.87f^{0.25}, \quad (19)$$

where $\sigma(f)$ is in radians.

Directional spectra are then estimated from the simulated time-series using kriging, adaptive kde2D and kde1D methods. Figure 12 shows an example of estimated directional spectrum in comparison to the reference one. Since the directional spectrum is the product of $S_{zz}(f)$ by $D(f, \theta)$ the same results on spectral moments may be obtained as in 1D case. The directional width $\sigma(f)$ seems the natural parameter that can be investigated for checking the quality of directional spectra estimates. The Figure 13 presents the result of directional width estimates with the three methods for 500 consecutive time steps in comparison of the reference directional width (19). Directional width obtained with kde1D method exhibits the most highest dispersion compared to the other methods. The mean difference between quantile 5% and quantile 95% is 5° against 3.5° for kriging and 3° for adaptive kde2D. These differences do not look significant, because the chosen reference directional width (West Africa climate) corresponds to narrow spectra; wide spectra might show greater differences since they are expected to be proportional to the directional width. The results however show that it is more likely to get accurate directional width estimates using kriging or (slightly better) adaptive kde2D.

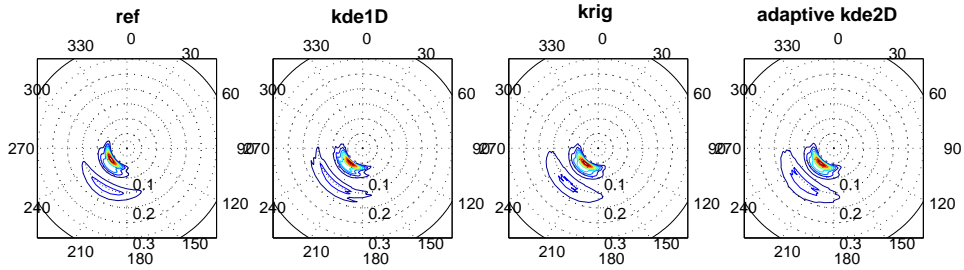


Fig. 12. Example of directional spectrum estimates

5 Conclusions

Two methods based on kriging and adaptive Kernel Density Estimator are proposed for both 1D and 2D ocean wave spectral estimation and interpolation. The main idea behind those two methods is the use of the temporal correlation structure of wave systems (wind sea and swell) to improve wave spectra estimation. Regarding 1D spectrum estimation, the first stage common to both methods is to calculate the periodograms (raw spectra) of recorded sea surface elevation time-series. The spectral density is then considered as a function of frequency and time and the next step is to treat the time-history of the log-periodogram as a random field and smooth it using either kriging or adaptive kde2D method. The methods are extended to 2D spectra estimation by using the same

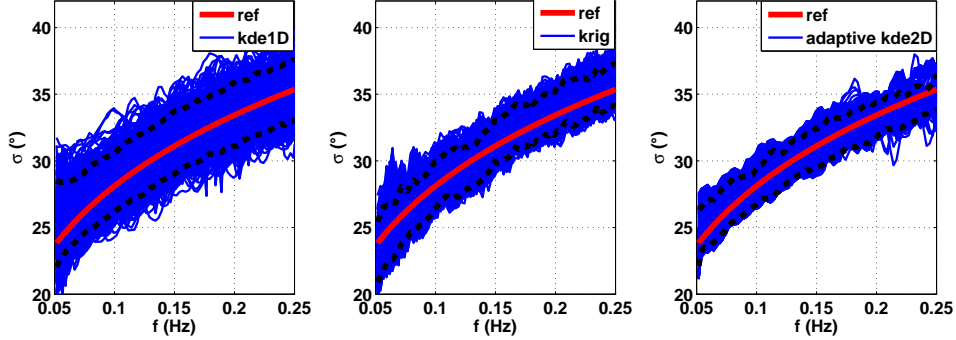


Fig. 13. Directional width obtained with the three methods; 5% and 95% fluctuation intervals are plotted in dashed black line. The red curve corresponds to the true value given by (19).

technique as for 1D spectra to estimate separately auto-densities and cross-densities, and then computing the directional distribution function with Fourier coefficients obtained from them.

The estimation methods were tested on synthetic yet field-similar data and the results were compared with the conventional kde1D method and a naive 2D averaging method. We have found that both Kriging and adaptive kde2D-based proposed methods provide better spectral estimates, i.e. a better representation of the spectra shape and better values of spectral moments compared to the kde1D method and a naive 2D averaging method. Regarding directional spectra, it is found that the proposed methods allow to better estimate the directional spreading parameter. It is also found that the adaptive kde2D-based method gives slightly better results than kriging. However, the adaptive kde2D method requires a prior model of time-evolution for peak frequencies of wind sea and swell systems. The kriging method has the advantage of not considering any prior model for the time-evolution of peak frequencies of wave systems. It is based on estimation of covariance matrices, the quality of which depends on the size of the data available.

The exploitation of the time dimension correlation structure in spectral estimation is also demonstrated to be a good tool for spectra interpolation over gaps when dealing with in-situ data, which often suffer from such missing data. With the proposed methods, we found that spectra can be properly interpolated up to 12h gap lengths, with the quality of interpolated spectra decreasing when gaps length increases.

Some improvements can still be brought to the proposed estimation methods. Especially, for adaptive kde2D, the use of wind data may be used for the improvement of the time-evolution model of peak frequencies of the wind sea. Regarding kriging, a choice of an appropriate covariance structure model rather than using the empirical one may improve the spectral estimates.

The proposed estimation methods can improve the tracking of wind sea and swell systems as much as the wave systems time evolution is taken into account in spectra estimation. Kriging might also be used when one accesses data from an array of buoys. It can similarly be applied to swell time-space tracking.

Acknowledgment

The data used in the study is the joint property of TOTAL E&P Angola and Esso Exploration Angola (Block 17) Ltd. / BP Exploration (Angola) Ltd. / Statoil Angola Block 17 A.S. / Norsk

Hydro Dezassete A.S. / SONANGOL, Sociedade Nacional de Combustveis de Angola, EP. The authors of the papers are grateful to them for the authorization to use the data for this paper.

References

- [1] J. Aarnes and H. Krogstad. Partitioning sequences for the dissection of directional ocean wave spectra: A review. *Part of ENVIWAVE Project*, 2001.
- [2] P. Ailliot, C. Maisondieu, and V. Monbet. Dynamical partitioning of directional ocean wave spectra. *Probabilistic Engineering Mechanics*, 33:95–102, 2013.
- [3] R. B. Blackman and J. W. Tukey. *The measurement of power spectra: from the point of view of communications engineering*, volume 190. Dover New York, 1959.
- [4] L. E. Borgman. Ocean wave simulation for engineering design. Technical report, DTIC Document, 1967.
- [5] A. W. Bowman. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71:353–360, 1984.
- [6] P. J. Brockwell and R. A. Davis. Time series: Theory and methods. *New York: Springer-Verlag. Christensen, R. (1991). Linear Models for Multivariate, Time Series and Spatial Data, New York: Springer-Verlag. Cliff, AD and Ord, JK (1974). "Space-time modeling with an application to regional forecasting," Transactions and Papers of the Institute of British Geographers, 66:119–128, 1991.*
- [7] N. Cressie. *Statistics for spatial data: Wiley series in probability and statistics*. 1993.
- [8] S. Elgar. Bias of effective degrees of freedom of a spectrum. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 113(1):77–82, 1987.
- [9] J. Fan and E. Kreutzberger. Automatic local smoothing for spectral density estimation. *Scandinavian Journal of Statistics*, 25(2):359–369, 1998.
- [10] G. Z. Forristal, K. Ewans, M. Olagnon, and M. Prevosto. The west africa swell project (wasp). In *Proceedings of the 32nd Int. Ocean, Offshore and Artic Eng. Conf., ASME. paper number 11264*, 2013.
- [11] T. W. Gerling. Partitioning sequences and arrays of directional ocean wave spectra into component wave systems. *Journal of atmospheric and Oceanic Technology*, 9(4):444–458, 1992.
- [12] B. Gjevik, H. Krogstad, A. Lygre, and O. Rygg. Long period swell wave events on the norwegian shelf. *Journal of physical oceanography*, 18(5):724–737, 1988.
- [13] Y. Goda. Numerical experiments on statistical variability of ocean waves. *Report of the Port and Harbour Research Institute*, 16(2):3–26, 1977.
- [14] Y. Goda. The observed joint distribution of periods and heights of sea waves. *Coastal Engineering Proceedings*, 1(16), 1978.
- [15] Y. Goda. *Random seas and design of maritime structures*. World Scientific, 2010.

- [16] J. Hannig and T. Lee. Kernel smoothing of periodograms under kullback–leibler discrepancy. *Signal processing*, 84(7):1255–1266, 2004.
- [17] J. L. Hanson and O. M. Phillips. Automated analysis of ocean surface directional wave spectra. *Journal of atmospheric and oceanic technology*, 18(2):277–293, 2001.
- [18] K. Hasselmann, T. P. Barnett, E. Bouws, H. Carlson, D. E. Cartwright, K. Enke, J. A. Ewing, H. Gienapp, D. E. Hasselmann, P. Kruseman, et al. Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap). 1973.
- [19] T. C. M. Lee. A simple span selector for periodogram smoothing. *Biometrika*, 84(4):965–969, 1997.
- [20] M. S . Longuet-Higgins, D. E. Cartwright, and N. D. Smith. Observations of the directional spectrum of sea waves using the motions of a floating buoy. 1963.
- [21] M. Olagnon, K. Ewans, G.Z. Forristall, and M. Prevosto. West africa swell spectral shapes. In *Proc. 32nd Int. Ocean, Offshore and Arctic Eng. Conf., OMAE 2013-11228*, 2013.
- [22] M. Olagnon and A. Robin. Numerical simulation of actual irregular wave properties. 1990.
- [23] Y. Pawitan and F. O’sullivan. Nonparametric spectral density estimation using penalized whittle likelihood. *Journal of the American Statistical Association*, 89(426):600–610, 1994.
- [24] O.M. Phillips. Spectral and statistical properties of the equilibrium range in wind-generated gravity waves. *Journal of Fluid Mechanics*, 156:505–531, 1985.
- [25] J. Portilla, F. J. Ocampo-Torres, and J. Monbaliu. Spectral partitioning and identification of wind sea and swell. *Journal of atmospheric and oceanic technology*, 26(1):107–122, 2009.
- [26] A. Schuster. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism*, 3(1):13–41, 1898.
- [27] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.
- [28] J. S. Simonoff. *Smoothing methods in statistics*. Springer, 1996.
- [29] M. L. Stein. *Interpolation of spatial data: some theory for kriging*. Springer, 1999.
- [30] P. Stoica and L. Randolph. *Introduction to spectral analysis*, volume 1. Prentice hall New Jersey:, 1997.
- [31] H. U. Sverdrup and W. H. Munk. Wind sea, and swell: theory of relations for forecasting. 1947.
- [32] G. Wahba. *Automatic Smoothing of the log periodogram*. University of Wisconsin, Department of Statistics, 1978.
- [33] M.P. Wand and M.C. Jones. Comparison of smoothing parameterizations in bivariate kernel density estimation. *Journal of the American Statistical Association*, 88(422):520–528, 1993.

- [34] P. Welch. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *Audio and Electroacoustics, IEEE Transactions on*, 15(2):70–73, 1967.
- [35] Y. Zhang, A. B. Baggeroer, and J. G. Bellingham. The total variance of a periodogram-based spectral estimate of a stochastic process with spectral uncertainty and its application to classifier design. *Signal Processing, IEEE Transactions on*, 53(12):4556–4567, 2005.